Improving on a systemic problem with the Extended Classifier System (XCS)

Charalambos Ioannides
Kerstin Eder, Geoff Barrett
Proposed Solution

[Diagram with labeled nodes and arrows]

- Pseudorandom Generator
- 0100110
- XCS
- 0100110
- fire_drill
- Test
- SIMULATOR
- DUV

[Table with columns: MAC, MMV, MSUM, MUL, OTHER, MREG0-3, SIMD_ST64-111, BIASES, COVERAGE]

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>0</th>
<th>0</th>
<th>1</th>
<th>1</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAC</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>SIMD_ST64</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BIASES</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>COVERAGE</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
### DV1 Function

<table>
<thead>
<tr>
<th>Input</th>
<th>0000</th>
<th>0001</th>
<th>0010</th>
<th>0011</th>
<th>0100</th>
<th>0101</th>
<th>0110</th>
<th>0111</th>
<th>1000</th>
<th>1001</th>
<th>1010</th>
<th>1011</th>
<th>1100</th>
<th>1101</th>
<th>1110</th>
<th>1111</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>11</td>
<td>17</td>
<td>19</td>
<td>19</td>
<td>19</td>
<td>19</td>
<td>19</td>
<td>19</td>
<td>19</td>
<td>19</td>
<td>19</td>
<td>19</td>
<td>19</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>5</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>11</td>
<td>17</td>
<td>19</td>
<td>19</td>
<td>19</td>
<td>19</td>
<td>19</td>
<td>19</td>
<td>19</td>
<td>19</td>
<td>19</td>
<td>19</td>
<td>19</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>11</td>
<td>17</td>
<td>19</td>
<td>19</td>
<td>19</td>
<td>19</td>
<td>19</td>
<td>19</td>
<td>19</td>
<td>19</td>
<td>19</td>
<td>19</td>
<td>19</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>15</td>
<td>31</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>23</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>14</td>
<td>30</td>
<td>22</td>
<td>22</td>
<td>22</td>
<td>22</td>
<td>22</td>
<td>22</td>
<td>22</td>
<td>22</td>
<td>22</td>
<td>22</td>
<td>22</td>
<td>22</td>
<td>22</td>
</tr>
</tbody>
</table>
How XCS works

1. Environment
2. Population [P]
   Classifier = condition : action = parameter(s)
3. Match set [M]
   Classifier
4. Covering
5. Action selection
6. Action set [A]
   Classifier
7. Action performed
8. Learning strategy
   Credit assignment
9. Update parameter(s)
10. Genetic algorithm

- Discovery component
- Performance component
- Reinforcement component
• Evolves rules that, given a problem, form a mapping between its input and output variables with the following properties:
  • Complete
  • Minimal (maximally general rules)
  • Accurate
• A fourth property talks about the disjunctiveness of rules
• Our previous work has experimentally disproved its validity
Problem Encountered

• Problem Description:
  • When a Boolean function requires overlapping classifiers, XCS has a problem retaining them in the Population [P]
  • Suboptimal Fitness is assigned to optimal rules which makes them prone to deletion
• Problem Source:
  1. Boolean Function’s structure
  2. Ternary Representation {0,1,)#
  3. XCS’ Fitness Sharing mechanism
DV1 – Overlap Example

Classifier 1: ###101#
Classifier 2: 1####10

Overlapping Region: 1##1010
Analysis of Problem

1. Discover the rules expected to be learnt by XCS
2. Examine these rules, establish why they should be part of the optimal population to be learnt and how they overlap with the rest
3. Discover the critical factors relating these optimal rules together
4. Provide solutions based on altering XCS in terms of these factors
Discovering the rules

- To do this we use the Quine-McCluskey Algorithm
  - Exact Logic Minimization Algorithm
  - Finds all Prime Implicants
  - Finds all Essential Prime Implicants
  - Finds one set of Minimal Prime Implicants
Prime Implicant Sets

* = distinguished 1-cell

Implicants:
0101
1110

Prime Implicants:
#101
#110
11#1
111#

Essential Prime Implicants:
#101
#110

MPI₁:
#101
#110
11#1

MPI₂:
#101
#110
111#
Solutions Proposed (1)

• Essentiality Assessment (T1)
  • If cl.exp>theta_{SUB} (experienced) and cl.error > \epsilon_0 (accurate) and it is a potentially essential prime implicant then it does not share its fitness with the rest of classifiers in [A]
  • A potentially essential classifier is one that covers at least one minterm, that no other equally general, experienced and accurate classifier does.
Solutions Proposed (2)

• Individualized Learning Rate (T2)

\[ F_{t+1} = F_t + \beta(F_{\text{new}} - F_t) \]

• Instead of a fixed learning rate (\(\beta\)), each classifier uses a different learning rate to update its stats based on its coverage in the search space (i.e. the generality of the classifier’s condition)

\[ \beta = \frac{1}{2^h} \]

where \(h\) is the number of (#) symbols in the classifier’s condition
## Results

<table>
<thead>
<tr>
<th></th>
<th>Orig.</th>
<th>Raw Acc.</th>
<th>T1</th>
<th>T2</th>
<th>T1+2</th>
<th>T2_F</th>
<th>T1+2_F</th>
</tr>
</thead>
<tbody>
<tr>
<td>%MPI</td>
<td>0.48</td>
<td>0.63</td>
<td>0.52</td>
<td>0.75</td>
<td>0.73</td>
<td>0.38</td>
<td>0.42</td>
</tr>
<tr>
<td>%EPI</td>
<td>0.51</td>
<td>0.63</td>
<td>0.53</td>
<td>0.80</td>
<td>0.74</td>
<td>0.41</td>
<td>0.44</td>
</tr>
<tr>
<td>%PI</td>
<td>0.45</td>
<td>0.60</td>
<td>0.48</td>
<td>0.71</td>
<td>0.70</td>
<td>0.33</td>
<td>0.36</td>
</tr>
<tr>
<td>AVG PERF</td>
<td>98.16</td>
<td>98.05</td>
<td>98.70</td>
<td>99.68</td>
<td>99.76</td>
<td>96.94</td>
<td>98.05</td>
</tr>
<tr>
<td>AVG ERR</td>
<td>101.72</td>
<td>91.66</td>
<td>95.09</td>
<td>34.21</td>
<td>39.01</td>
<td>162.18</td>
<td>151.32</td>
</tr>
<tr>
<td>AVG POP</td>
<td>129.65</td>
<td>176.10</td>
<td>129.94</td>
<td>219.10</td>
<td>226.67</td>
<td>161.92</td>
<td>168.65</td>
</tr>
<tr>
<td>AVG MPI</td>
<td>27.14</td>
<td>36.19</td>
<td>29.06</td>
<td>44.27</td>
<td>44.99</td>
<td>20.05</td>
<td>23.65</td>
</tr>
<tr>
<td>AVG EPI</td>
<td>22.20</td>
<td>28.22</td>
<td>22.89</td>
<td>35.54</td>
<td>35.25</td>
<td>16.99</td>
<td>19.48</td>
</tr>
<tr>
<td>AVG PI</td>
<td>34.38</td>
<td>47.04</td>
<td>37.04</td>
<td>57.80</td>
<td>59.20</td>
<td>23.37</td>
<td>28.18</td>
</tr>
</tbody>
</table>
Essentiality Assessment (T1)
Individualized $\beta$ (T2)
Future Research

- Investigate further the individualized learning rate success
- Solve the overlapping classifier problem with XCS
- Use %EPI as an essential metric to judge on XCS population state metric
- Use XCS in larger DV problems and using different representations
THANK YOU

Any Questions?
Insights from Improvements

- **Essentiality Assessment:**
  - Essential Prime ‘hunt’ is improved by searching for specific properties
  - Removing fitness competition between them

- **Individualized Learning Rate:**
  - Favours learning from Specific to General rules (enhancing PI rule learning)
  - Prolonged “memory” on classifier stats
Proposed Solution
Interacting Pressures

# Prime Implicant Retention

<table>
<thead>
<tr>
<th></th>
<th>Orig.</th>
<th>Raw Acc.</th>
<th>T1</th>
<th>T2</th>
<th>T1+2</th>
<th>T2_F</th>
<th>T1+2_F</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MPI/POP</strong></td>
<td>0.23</td>
<td>0.21</td>
<td>0.24</td>
<td>0.20</td>
<td>0.20</td>
<td>0.15</td>
<td>0.15</td>
</tr>
<tr>
<td><strong>EPI/POP</strong></td>
<td>0.18</td>
<td>0.16</td>
<td>0.19</td>
<td>0.17</td>
<td>0.16</td>
<td>0.12</td>
<td>0.13</td>
</tr>
<tr>
<td><strong>PI/POP</strong></td>
<td>0.30</td>
<td>0.27</td>
<td>0.31</td>
<td>0.27</td>
<td>0.26</td>
<td>0.17</td>
<td>0.18</td>
</tr>
</tbody>
</table>